

Mapping for Prevention: GIS Models for Directing Childhood Lead Poisoning Prevention Programs

Marie Lynn Miranda, Dana C. Dolinoy, and M. Alicia Overstreet

Nicholas School of the Environment and Earth Sciences, Duke University, Durham, North Carolina, USA

Environmental threats to children's health—especially low-level lead exposure—are complex and multifaceted; consequently, mitigation of these threats has proven costly and insufficient and has produced economic and racial disparities in exposure among populations. Policy makers, public health officials, child advocates, and others currently lack the appropriate infrastructure to evaluate children's risk and exposure potential across a broad range of risks. Unable to identify where the highest risk of exposure occurs, children's environmental health programs remain mitigative instead of preventive. In this article we use geographic information system spatial analysis of data from blood lead screening, county tax assessors, and the U.S. Census to predict statistically based lead exposure risk levels mapped at the individual tax parcel unit in six counties in North Carolina. The resulting model uses weighted risk factors to spatially locate modeled exposure zones, thus highlighting critical areas for targeted intervention. The methods presented here hold promise for application and extension to the other 94 North Carolina counties and nationally, as well as to other environmental health risks. *Key words:* children's health, environmental justice, geographic information system, lead. *Environ Health Perspect* 110:947–953 (2002). [Online 13 August 2002]

<http://ehpnet1.niehs.nih.gov/docs/2002/110p947-953miranda/abstract.html>

Until now, research studies on childhood lead exposure risk have identified risk factors but not considered the relative weight for each factor (1,2), considered relative weights but not linked analysis to geographic location (3), or linked analysis to geographic location at highly aggregated levels (i.e., census block or U.S. Postal Service ZIP code) but not considered relative weights (4–9). In this study we estimate and apply relative weights for risk factors at a very high geographic resolution: the individual tax parcel unit level. In so doing, this study provides a strong basis for shifting from mitigative to preventive intervention programs aimed at protecting children from the adverse effects of lead.

Medical professionals have long recognized severe lead poisoning as a debilitating disease. Since the late 1970s, however, mounting research has shown that lead also causes asymptomatic effects at levels far below thresholds previously considered safe (10–16). The adverse effects of lead, including learning and behavioral disorders (e.g., attention deficit disorder and attention deficit hyperactivity disorder), hearing impairment, decreased intelligence quotient, and decreased attention span, are particularly harmful in children and often become apparent during puberty—long after exposure has caused irreversible impacts (10–16). Thus, the Centers for Disease Control and Prevention (CDC) have lowered incrementally the threshold for lead levels considered dangerous in children by 88%—from 60 to 10 $\mu\text{g}/\text{dL}$ —in the last 40 years. Furthermore, a new body of literature suggests that cognitive deficits may occur

at blood lead levels as low as 5 $\mu\text{g}/\text{dL}$ (17,18). For example, a recent study observed an inverse relationship between blood lead concentrations below 5 $\mu\text{g}/\text{dL}$ and scores on reading and mathematics tests (17).

Despite substantial gains from the elimination of leaded gasoline, nearly one million U.S. children still have blood lead levels above the current CDC threshold of 10 $\mu\text{g}/\text{dL}$ (19). Current exposures result primarily from environmental sources of lead incorporated into infrastructure, including paint, water systems, and soil. These intransigent sources of lead are difficult and costly to abate; thus, protecting children from lead exposure remains a daunting task. Because even the most capable doctors cannot easily diagnose low-level lead exposure, screening for lead in high-risk populations is critical to eradicating this disease. In addition, shifting to preventive rather than mitigative approaches requires characterization of the housing stock for exposure risk and abatement of sources of biologically available lead.

Exposure to lead-based paint is the leading cause of childhood lead poisoning today (20). Fifty million U.S. homes still contain lead-based paint, with lead concentrations anywhere from 1% to 50% by dry weight (21). Because of the expense of upkeep, lead-based paint found in older, low-income housing runs the greatest risk of being in poor condition. Young children easily ingest chips of lead-based paint, which tastes sweet. Household dust containing lead particles can be more dangerous than paint chips, because smaller particles are more easily absorbed by both the gastrointestinal and pulmonary

tracts. By creating lead-contaminated dust from already-existing sources, including those that were previously undisturbed, both household renovation and attempts to remove lead-based paint can increase levels of biologically available lead in the home (21).

A January 1999 report issued by the U.S. General Accounting Office (GAO) revealed that children in or targeted by federal health care programs [e.g., Medicaid, the Women, Infants, and Children (WIC) program, and community health centers] exhibit elevated blood lead levels at nearly five times the rate of other children (22). Despite federal policies requiring blood lead screening of these children, less than 20% served by federal health care programs are actually screened (22).

Over the 4-year period from 1995 through 1998, 373,619 children were screened for lead in North Carolina (Table 1) (23). Of these, 4.8% had blood lead levels at or above the CDC's threshold of 10 $\mu\text{g}/\text{dL}$. Minority children exhibited the highest prevalence rates, with African Americans at 7.2%, Native Americans at 6.2%, and Hispanics at between 5% and 6%. Although computation of blood lead data is complicated, we can reasonably infer that a high percentage of children in the "Other" category are Hispanic. Prevalence among white children was 3.1%. Prevalence among children receiving WIC assistance was 6.8%, compared with 3.8% for non-WIC children. In 1998, 25% of children 1–2 years old underwent screening in North Carolina, with 3.6% exhibiting elevated blood lead levels. Consistent with the findings of the GAO

Address correspondence to M. L. Miranda, Nicholas School of the Environment and Earth Sciences, Duke University, Box 90328, Levine Science Research Center Room A134, Durham, NC 27708 USA. Telephone: (919) 613-8023. Fax: (919) 684-8741. E-mail: mmiranda@duke.edu

We thank S. Buchanan, M. Lavine, E. Norman, W. Thomann, T. Ward, and two anonymous referees for their review and comment on this work. Able research assistance was provided by C. Bradshaw, K. Engleman, and D. Jones.

This research was made possible by funding from the Centers for Disease Control and Prevention, the National Institute of Environmental Health Sciences, the Mary Duke Biddle Trent Semans Foundation, and the Duke University Office of the Senior Vice President for Research Administration and Policy.

Received 31 August 2001; accepted 21 February 2002.

report (22), 42.9% of children receiving Medicaid and 37.6% of children receiving WIC assistance in North Carolina had undergone screening in 1997 (23).

Literature Review

In examining the factors that influence the risk of lead exposure and uptake in children, it is important to recognize the interrelationships among risk factors. Age of housing, urban/rural status, race/ethnicity, socioeconomic status, and nutritional status all relate to and influence one another, especially regarding childhood lead poisoning. However, studies disagree on which factors are the most significant predictors of lead poisoning (7).

Data from phase 2 of the third National Health and Nutrition Examination Survey (NHANES III) (24) reveal a relatively high prevalence of elevated blood lead levels in children who live in housing built before 1973, as well as in children who live in metropolitan areas with populations greater than one million. Lanphear et al. (7) associate high population density, older housing, renter-occupied housing, and lower housing value with childhood lead exposure in Monroe County, New York. In contrast to the Northeast and Midwest, in North Carolina older rural housing contains

more lead-based paint than urban housing and poses a greater risk of lead poisoning (2). As is true for much of the southeastern United States, urban centers in North Carolina (e.g., Charlotte, Greensboro, and Raleigh–Durham) experienced their major growth phase—in terms of both people and new housing stock—in the 1980s onward, after lead was banned from use in paint.

Identification of the socioeconomic and racial/ethnicity status of residential neighborhoods can help determine a child's risk level for low-level lead poisoning. Sargent et al. (3) evaluated risk factors for childhood lead exposure in Massachusetts and identified several significant independent associations, including percentage of single-parent households, median income, percentage African American, percentage of children in poverty, percentage of renter-occupied housing, median age of housing, and blood lead screening rates (3). According to phase 2 of NHANES III, the prevalence of elevated blood lead levels for children from low-income families was 8.0%,

eight times higher than that for children from high-income families (19). The prevalence among non-Hispanic black children was 11.2%, almost five times higher than that among non-Hispanic white children of the same age (2.3%) (19). The prevalence among Mexican-American children was 4.0%, nearly twice that of white children. Additionally, non-Hispanic black race/ethnicity is an independent predictor of elevated blood lead levels for children between 1 and 5 years old (24).

Although race has been demonstrated as an independent predictor for elevated blood lead levels, it is unclear whether race serves as a proxy for other conditions that may pose risk (e.g., racial segregation) or whether a race-based physiologic difference in uptake exists. For example, according to NHANES II, one reason black children are at higher risk of lead poisoning may be that blacks have lower intakes of dietary calcium than do whites, a finding that has been corroborated by several studies (25). Reasons for lower dietary calcium in blacks include lactose

Table 1. Estimated prevalence of elevated blood lead levels among North Carolina children screened for lead poisoning, 1995–1998.

Variable	Subcategory	Prevalence (%)
Race ^a	Black	7.2
	White	3.1
	Native American	6.2
	Other	5.9
Residence	Rural county	5.2
	Urban county	3.8
Income ^b	WIC	6.8
	Non-WIC	3.8
Overall		4.8

^aThere are missing values for these variables. ^bAvailable only for 1995 and 1996. Data from NCCLPPP (23).

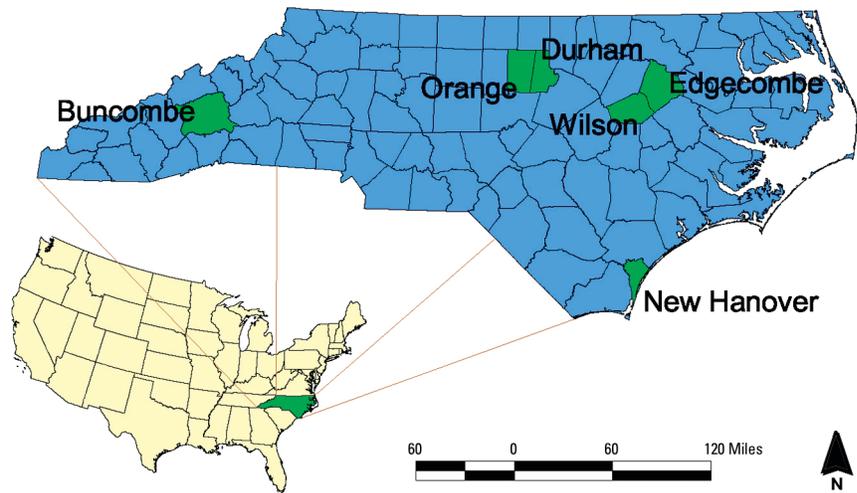


Figure 1. Study area counties.

Census geographic units

- Blocks
- Block groups
- Tracts

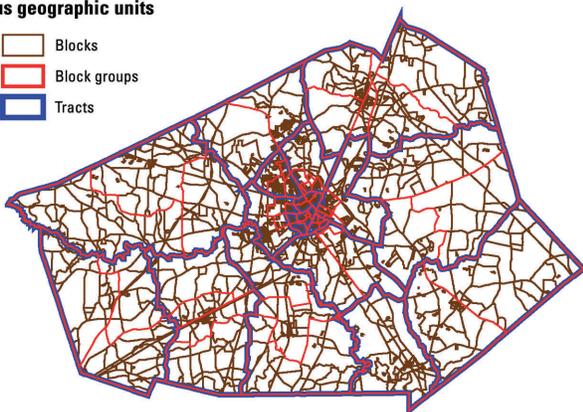


Figure 2. Geographic resolution of U.S. Census data, Wilson County, North Carolina.

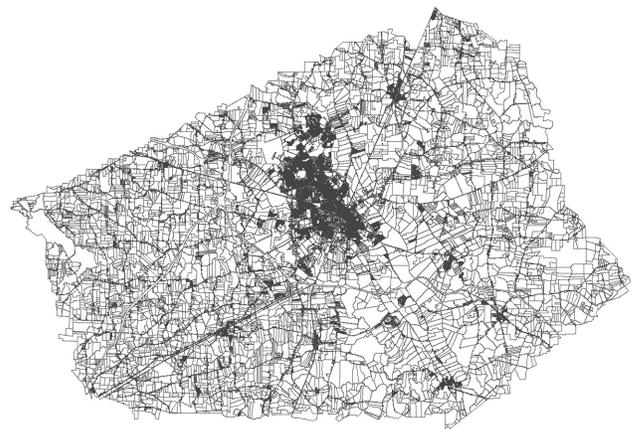


Figure 3. Tax parcel units for Wilson County, North Carolina.

intolerance, cultural unfamiliarity with drinking milk, poverty, restricted access to market, and limited food storage facilities (25).

Several recent studies have used spatial analysis and geographic information system (GIS) technology to compare the spatial distribution of blood lead levels with identified risk factors for exposures (4–9). These studies were implemented at the census tract, block group, block, and/or U.S. Postal Service ZIP code level of resolution and build upon previous work that developed guidelines for using GIS technology in environmental epidemiology research and lead exposure analysis (26,27).

Thus, considerable knowledge exists about the risk factors for childhood lead exposure. These include age, socioeconomic status, race/ethnicity, nutritional status, and age and urban/rural status of housing. Yet this knowledge has not translated successfully into proactive and preventive strategies to eradicate the

threat to children. One reason may be that the scientific literature fails to characterize adequately the importance of each factor relative to the others. Furthermore, previous research studies have not fully analyzed geographic location as a predictor for low-level lead poisoning. This study extends previous work by estimating exposure risk across a variety of risk factors at a very fine geographic resolution.

Methods

The flexibility and comprehensiveness of GIS technology and spatial analysis allow the integration of multifactorial components in an aggregate risk model. The key to spatial analysis is that most data contain a geographic component that can be tied to a specific location, such as a country, state, county, ZIP code, census block, or specific address. Geographic coding allows users to explore and overlay data by location, revealing relationships that are not readily apparent in traditional spreadsheet

and/or statistical packages. Additionally, GIS technology has specific capabilities that allow users to produce clear and accessible maps and data reports that can serve as powerful community outreach tools.

Using GIS technology as well as statistical analysis, we have developed a predictive exposure model for low-level childhood lead poisoning for six North Carolina counties. Figure 1 shows the location of the six study counties. The counties represent four distinct geographic sectors: Buncombe in the western portion of the state, Durham and Orange in the central piedmont, Wilson and Edgecombe in the eastern coastal plain, and New Hanover on the southeast coast. Including study counties from the mountains, piedmont, coastal plain, and coast allows for comparisons across regional, topographic, economic, and cultural zones. These variations are important for characterizing risk across spatial dimensions.

Data

We used U.S. Census demographic data, county tax assessor data, and North Carolina blood lead screening data to construct the lead exposure model.

U.S. Census data. Census demographic information is available in three different geographic scales: tracts, block groups, and blocks. Tracts designate the largest geographic areas. The most detailed and focused information is contained in blocks. Blocks are also combined into block groups, an intermediate category. Our GIS county projects contain attribute themes for median household income, percentage of children in poverty, percentage of persons in poverty, percentage of renter-occupied households, percentage of single-parent households, percentage of African Americans, and number of Hispanics from the 1990 Census (29). Census variables can be custom divided to target specific demographic and socioeconomic groups. Figure 2 maps the census tract, block group, and block outlines for Wilson County, North Carolina. Most previous GIS studies of environmental health issues have been applied at this level of geographic resolution.

In addition to demographic data, county models contain 1995 topologically integrated geographic encoding and referencing (TIGER) census street data (29). The TIGER data provide information on street names, locations, and address numbers and are extremely useful for converting the outcomes from the research project into direct service public health programs.

Tax assessor data. County tax assessor offices track a wide variety of information on individual tax parcel units, all of which is publicly available. We focus on residential (vs. undeveloped or commercial) tax parcel units. Residential tax parcel units typically

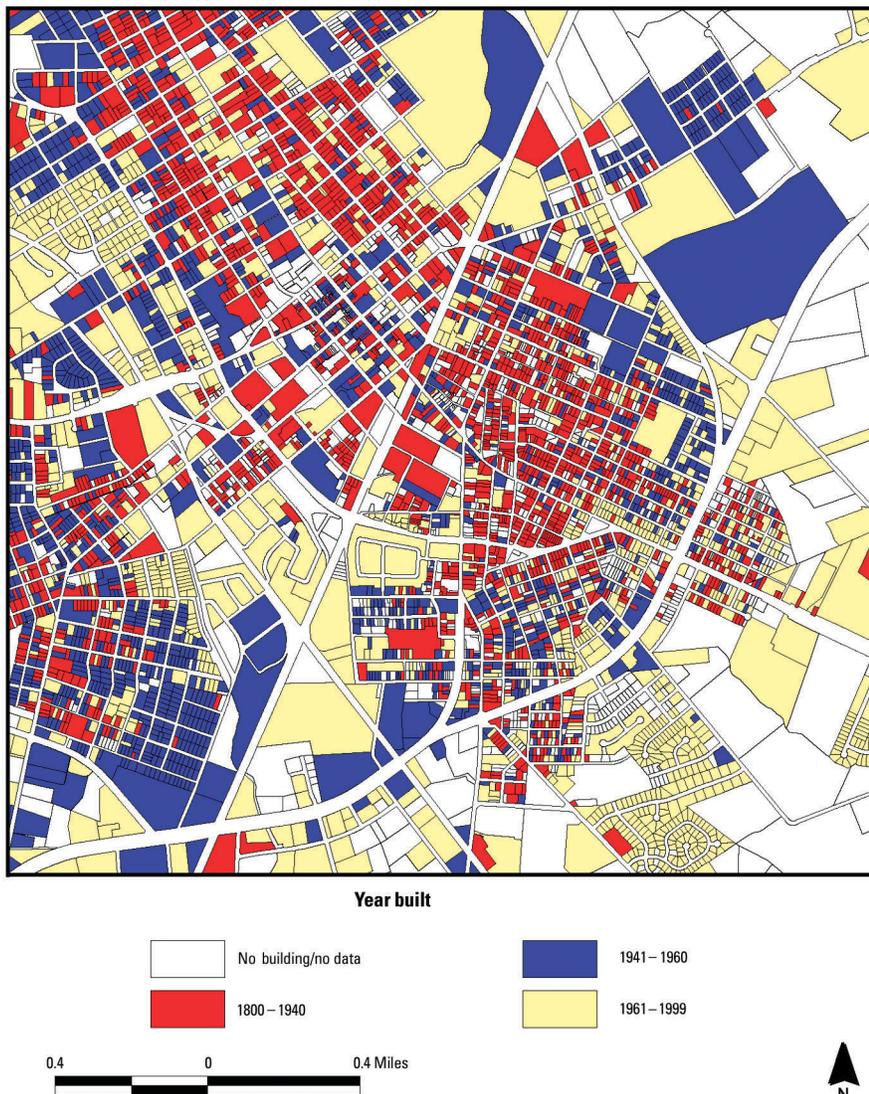


Figure 4. Age of housing mapped at the tax parcel unit level in a Wilson County, North Carolina, neighborhood.

encompass a housing structure (either single or multifamily) and its yard. A residential tax parcel unit may be owned by a single individual or by a group; it may be unoccupied, occupied by the owner, or occupied by nonowners—either renters or otherwise. The presence of specific county tax variables depends on the sophistication of the data monitored and targeted by the individual county tax assessor office. Each of our county models contains tax assessor parcel information about year of construction. Other variables of interest include assessed tax value, parcel unique identifier code, building class, date remodeled (if any), construction type, zoning codes, use codes, owner address, physical address, tax district, and renter/owner-occupied status. Each of these variables speaks to the general state of maintenance and repair of the housing unit.

Figure 3 shows a map of Wilson County, North Carolina, with each of the 34,928 individual tax parcel units outlined in black (28). The dense black area in the middle of the county is the town of Wilson. Figure 4 zooms in on a section of the town of Wilson and maps age of housing in three categories to show the level of detail made possible by tax parcel level analysis.

Comparing Figures 3 and 4 with Figure 2 illustrates the enhanced analytic potential associated with the high geographic resolution used in this study.

Blood lead screening data. Through a negotiated confidentiality agreement with the North Carolina Childhood Lead Poisoning Prevention Program (NCCLPPP), each model incorporates childhood lead screening data for children born and screened between 1994 and 1999 (23). The screening data for each county are geocoded to the individual tax parcel unit (vs. street block) using the

household-level tax assessor databases. The screening data include the child's name, birth date, test date, blood lead level, and address. We also consolidated duplicate child screens from the same residence. We take a conservative or protective approach (in terms of identifying biologically available lead) by retaining entries with the highest blood lead level, which is consistent with Lanphear et al. (7). Match rates range from 53% to 86% across counties. We were not able to geocode children who did not list an address or listed addresses that contain post office boxes or were incomplete. In addition, the state database included multiple examples of children whose screening results were assigned to the wrong county. We deleted these observations from our analysis. Overall, our match rates compare favorably with previous studies. Table 2 presents target population and screening rates for children 0–2 years old (1995–1998) as well as geocoding match rates for children 0–6 years old (1995–1999) in each of the six study counties. The overall match rate indicates geocoding percentage matched before elimination of incomplete or post office box addresses. The trimmed match rate represents percentage matched after deletion of incomplete and post office box addresses. A previous study using address geocoding reported match rates of 20% in rural counties to 98% in urban counties (26). Lower match rates in less urban counties can be attributed to the difficulty in geocoding rural route addresses.

Previous GIS studies that obtained higher match rates geocode screening data to a street grid rather than to an individual tax parcel unit. These analyses typically use TIGER census street data that incorporate street location, street type, and address range. In comparing street geocoding using TIGER data with parcel

geocoding using tax assessor data, we determined that street geocoding often locates general house vicinity but rarely pinpoints the exact housing unit. Conversely, parcel geocoding locates children within the exact residential unit. As Figure 4 demonstrates, age of housing, an important risk factor, can vary substantially within the same block. Thus, geocoding to the tax parcel level provides a better basis for conducting statistical analysis.

Spatial and Statistical Analysis

We combined tax assessor, U.S. Census demographic, and North Carolina blood lead screening data into one spatial overlay theme. Although each of these data sets started out as a unique entity with a specific geographic resolution, they shared a common geographic spine. GIS software allowed us to combine the separate data sets into one large database, based on common geographic location. By integrating data, we were able to perform statistical analysis on all data layers together. The lead screening data served as the dependent variable and were used to calibrate the relative weights that should be assigned to each of the risk factors. We applied multivariate statistical analysis to 11,523 observations geocoded to the individual tax parcel unit.

On the basis of the existing literature on risk factors for lead exposure, we analyzed the relationship between observed blood lead levels geocoded to the individual tax parcel unit and age of housing, median income, percentage renter occupied, percentage of persons in poverty, percentage of children in poverty, percentage of one-parent households, and percentage African American as well as indicator variables for each of the six counties. Hispanics represent the fastest-growing subpopulation in North Carolina. The 1990 census data undercounted Hispanics, and that population has grown substantially in the past ten 10 years. We chose not to include number of Hispanics in our analyses because of the widely recognized poor quality of this these data in North Carolina. With the release of 2000 Census data, we hope to improve upon this portion of the

Table 2. Target population and screening rates for children 0–2 years old, 1995–1998, and GIS geocoding match rates for children 0–6 years old, 1995–1999.

Region	Target population ^a	Percentage screened ^a	Overall geocode match rate ^b	Geocode match rate for trimmed data ^b
Buncombe Co.	18,229	18.3	50.4	57.8
Durham Co.	24,196	21.7	60.7	66.7
Edgecombe Co.	7,049	56.0	47.2	53.4
New Hanover Co.	14,617	30.7	67.6	70.6
Orange Co.	8,997	16.8	72.1	86.5
Wilson Co.	8,014	12.0	63.3	73.8
North Carolina total	823,040	23.6	—	—

^aSource: NCCLPPP (23). ^bCalculated within GIS projects.

Table 3. Summary statistics on data used in the statistical analysis.

Explanatory variable	Data source	Geographic unit	Min/max	Mean	Median
Year built	County tax assessor office	Individual tax parcel	1787/1999	1964	1971
Black (%)	1990 U.S. Census	Block	0/100	35.90	13.59
Single parent (%)	1990 U.S. Census	Block	0/100	13.4	8.33
Renter-occupied (%)	1990 U.S. Census	Block	0/100	36.9	26.09
Median income (\$)	1990 U.S. Census	Block group	4,999/98,406	27,959	27,500
Children in poverty (%)	1990 U.S. Census	Block group	0/100	19.6	12.08

Abbreviations: Max, maximum; min, minimum.

Table 4. Results of the multivariate statistical analysis.

Variable	Coefficient	SD	p-Value
Year built	-0.044	0.00026	0.00
Median income	-4.42 × 10 ⁻⁶	6.84 × 10 ⁻⁷	0.00
Percentage African American	0.002	0.00021	0.00
Buncombe	9.85	0.5025	0.00
Durham	9.81	0.5017	0.00
Edgecombe	10.10	0.5021	0.00
New Hanover	9.92	0.5039	0.00
Orange	9.93	0.5042	0.00
Wilson	10.30	0.5044	0.00

F(9,11514) = 4816.12; adjusted R² = 0.7900; root mean square error = 0.61995

analysis. Table 3 lists each of the explanatory variables explored, as well as the data source, some descriptive statistics, and the geographic unit of analysis at which the data are coded.

We first examined the data using general additive models to search for the importance of nonlinear and county-specific effects. These analyses did not demonstrate any reason to favor nonlinear over linear models, although county-specific effects were noted. These analyses were, however, characterized by long-tailed residuals. Therefore, we estimated

log-linear models, where the dependent variable was given by $\ln[\max(\text{blood lead level}, 1)]$. The resulting models were characterized by well-behaved (i.e., Gaussian) residuals. Analysis of variance (ANOVA) analysis suggested that percentage of children in poverty, percentage of one-parent households, and percentage of renter-occupied housing did not add explanatory power, so they were dropped from model estimation. Percentage of persons in poverty and median income were both statistically significant variables if included indi-

vidually, with the latter appearing to be a better explanatory variable. As a result, we dropped percentage of persons in poverty from model estimation. We also examined a wide variety of interactive effects among variables, none of which appeared significant.

Once the general additive models approach failed to demonstrate nonlinear effects, we switched to log-linear ordinary least squares analysis with robust standard errors. Table 4 shows the statistical model we eventually used to construct the exposure risk indices.

All of the significant variables had the expected sign [e.g., the higher the median income, the lower the blood lead level (BLL); the older the home, the higher the blood lead level]. The coefficients were subsequently imported into the county GIS projects to construct a risk index value for each residential tax parcel unit in the county for which full data exist, according to the following six equations:

Buncombe:

$$\begin{aligned} \text{Estimated } \ln(\text{BLL}) &= 9.85 + (-0.0044 \times \text{year built}) \\ &+ (-4.42 \times 10^{-6} \times \text{median income}) \\ &+ (0.002 \times \text{percent African American}) \end{aligned}$$

Durham:

$$\begin{aligned} \text{Estimated } \ln(\text{BLL}) &= 9.81 + (-0.0044 \times \text{year built}) \\ &+ (-4.42 \times 10^{-6} \times \text{median income}) \\ &+ (0.002 \times \text{percent African American}) \end{aligned}$$

Edgecombe:

$$\begin{aligned} \text{Estimated } \ln(\text{BLL}) &= 10.10 + (-0.0044 \times \text{year built}) \\ &+ (-4.42 \times 10^{-6} \times \text{median income}) \\ &+ (0.002 \times \text{percent African American}) \end{aligned}$$

New Hanover:

$$\begin{aligned} \text{Estimated } \ln(\text{BLL}) &= 9.92 + (-0.0044 \times \text{year built}) \\ &+ (-4.42 \times 10^{-6} \times \text{median income}) \\ &+ (0.002 \times \text{percent African American}) \end{aligned}$$

Orange:

$$\begin{aligned} \text{Estimated } \ln(\text{BLL}) &= 9.93 + (-0.0044 \times \text{year built}) \\ &+ (-4.42 \times 10^{-6} \times \text{median income}) \\ &+ (0.002 \times \text{percent African American}) \end{aligned}$$

Wilson:

$$\begin{aligned} \text{Estimated } \ln(\text{BLL}) &= 10.30 + (-0.0044 \times \text{year built}) \\ &+ (-4.42 \times 10^{-6} \times \text{median income}) \\ &+ (0.002 \times \text{percent African American}). \end{aligned}$$

Previous GIS studies of childhood lead exposure encountered problems with spatial autocorrelation (5). At least two considerations make spatial autocorrelation problematic in childhood lead exposure analysis. To the

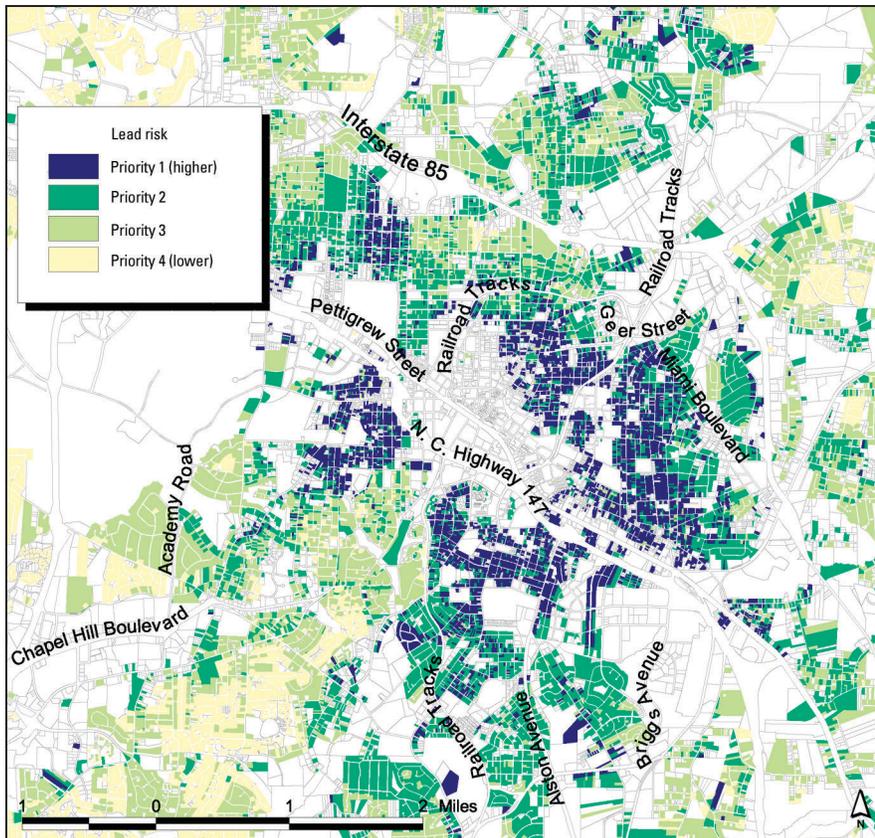


Figure 5. Lead risk priorities mapped at the tax parcel unit level for the city of Durham, North Carolina.

Table 5. Implications of model results for housing intervention programs.

County	Priority 1 (%)	Priority 2 (%)	Priority 3 (%)	Priority 4 (%)
Buncombe				
Housing stock	5	25	30	40
Elevated blood lead levels	30	25	20	25
Durham				
Housing stock	10	20	25	45
Elevated blood lead levels	35	30	30	5
Edgecombe				
Housing stock	10	25	35	25
Elevated blood lead levels	35	30	30	5
New Hanover				
Housing stock	5	15	35	45
Elevated blood lead levels	35	20	35	10
Orange				
Housing stock	10	20	30	40
Elevated blood lead levels	20	40	25	15
Wilson				
Housing stock	10	15	35	40
Elevated blood lead levels	40	30	20	5

extent that houses in neighborhoods or areas tend to be built at the same time and that neighbors tend to share common demographics, we may expect spatial autocorrelation problems within geographically based analyses. However, the very high geographic resolution at which our study was undertaken means that we have age of housing available at the tax parcel unit level. We tested whether the inclusion of age of housing at the tax parcel unit level is sufficient to eliminate problems with spatial autocorrelation in our model estimation. First, we plotted an empirical variogram of residuals against distance between observations (using the latitude and longitude measures available within the GIS). This variogram was flat. Second, an ANOVA comparison of models with and without spatial correlation is statistically insignificant. For these two reasons, we concluded that we did not need to include corrections for spatial autocorrelation in our model estimation.

Discussion

As a result of the statistical analysis, we created priority themes consisting of household-level maps coded by the lead exposure risk index. The county models contain priority themes with four categories. The categories are based on natural break statistical analysis. The basis for categorization, as well as the number of risk priority categories, is in some sense arbitrary. The flexibility of the GIS modeling approach allows for the construction of risk categories as they are useful to specific problems. For illustration, we chose four categories based on natural break statistical analysis given how counties are likely to use the GIS models in shaping preventive intervention programs. Alternative formulations using a different number of categories determined by standard deviates, quantiles, or other means are easy to implement with GIS techniques.

Presented below is a sample priority mapping drawn from the Durham County GIS project to demonstrate the usefulness of spatial analysis in identifying children at high risk for exposure to lead. Figure 5 depicts the priority categories for residences in the city of Durham, Durham County, North Carolina. Dark blue areas represent priority 1 parcels, which are predicted most likely to contain lead-based paint hazards. Priority 2 and 3 parcels are colored medium and light green and are less likely than priority 1 parcels to contain lead-based paint hazards. Yellow represents priority 4 parcels, which are least likely to contain lead-based paint hazards. White areas represent commercial or industrial properties. Based on this analysis, the corridors along Highway 147 represent areas that the Durham County Health Department may wish to target for lead abatement, public education, and community outreach efforts.

Compared with southern and northeastern Durham, central Durham is depicted with a high concentration of dark blue (priority 1) and green (priority 2 and 3) parcels.

Table 5 provides a sense of how the risk model might allow county health departments and community organizations to use scarce resources more effectively. It shows the percentage of the housing stock included in the priority 1–4 risk categories for the study counties. This analysis indicates that by focusing on 30% of the housing stock in Durham County, for example, intervention programs could address 70% of the estimated elevated blood lead levels. Table 5 provides analogous statistics for the other five study counties.

The detail provided by the GIS maps also allows for block- or even house-level planning for intervention programs. Using city marketing directories, state licensing agencies, and Internet searches, we created community databases that spatially locate businesses and institutions where children and parents tend to spend time, including schools, physicians' offices, churches, recreation facilities, and day care centers. In addition, we used county tax assessor data to spatially locate parks, playgrounds, swimming pools, and other public gathering grounds. Figure 6 provides a detailed example of how community databases are uploaded into GIS projects as point themes.

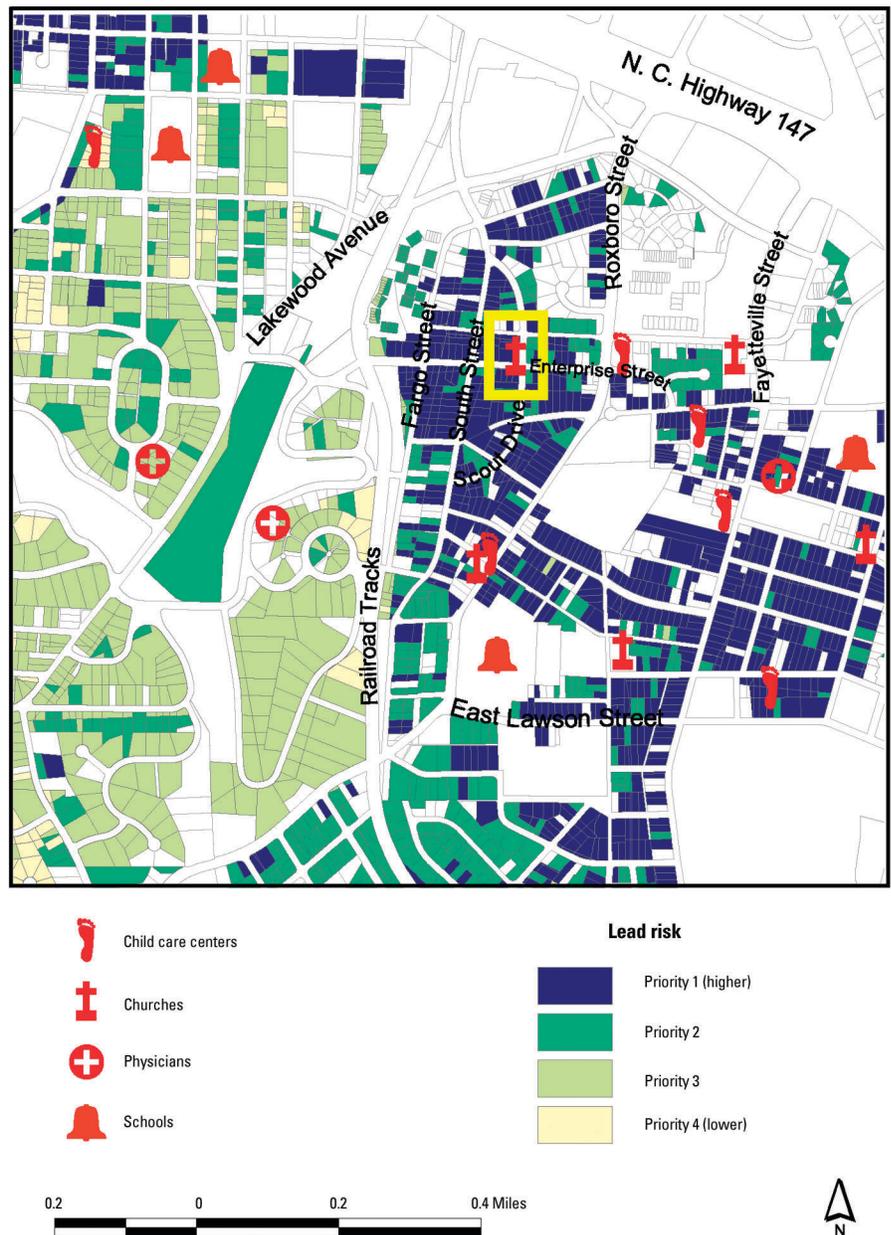


Figure 6. Community databases mapped at the tax parcel unit level for a Durham, North Carolina, neighborhood.

Figure 6 depicts the lead risk priority themes overlay for a Durham County neighborhood. The color coding for the priority categories is the same as for Figure 5. The red symbols indicate businesses and institutions where children and families tend to spend time. The local health department might, for example, sponsor a lead poisoning awareness health fair at the church located at the intersection of Scout Drive and Enterprise Street (outlined in yellow in Figure 6), an area characterized by a heavy concentration of priority 1 parcels.

The data-rich spatial analysis projects include a mechanism for personalized contact (using only publicly available data) with every homeowner in the study counties. Using the models in conjunction with city marketing directories allows for direct contact with most current residents (which makes tenants a reachable target group), as well as business owners and community leaders (e.g., the community database on churches includes the name and address of the pastor/priest/minister of the church).

Conclusions and Directions for Future Research

Policy makers, public health officials, child advocates, and others currently lack the appropriate infrastructure to evaluate children's potential exposure to lead across a broad range of risks. Unable to identify where the highest risk of exposure occurs, children's environmental health programs remain mitigative instead of preventive. Thus, children must first become sick before they can be protected. In this article we describe a predictive model of childhood lead exposure risk specified at the individual tax parcel unit level.

Although the model represents an important innovation over previous GIS-based work addressing childhood lead exposure risks, several limitations are important to note. The model development and statistical analysis rely on county tax assessor and North Carolina screening data quality. For example, parcel geocoding match rates depend on address accuracy reported by both the tax assessor and the State of North Carolina. Current models use 1990 Census data; as the 2000 Census data become available in GIS format, we will update our models. In addition, the 2000 screening data for North Carolina indicate higher screening rates, both statewide and within individual counties. As these data are released and GIS software evolves, we will incorporate these enhancements. In subsequent analyses, we also hope

to incorporate assessed tax value at the tax parcel unit level as a proxy for economic demographic characteristics. The approach described in this article will not capture non-housing-related aspects of lead exposure risk, such as cultural sources of lead (from traditional medicines or cosmetics) or hobby-related lead use (stained glass or fishing weights).

Besides updating models with current tax assessor and 2000 Census data, we are collecting environmental samples from homes in the study counties in order to validate and calibrate the childhood lead risk models. Environmental samples include at least 10 X-ray fluorescence readings, two dust wipe samples, and a composite soil sample to distinguish between the presence of lead and the presence of biologically available lead. This effort will further strengthen the analytic power of the models.

The model described in this article enables individuals and communities to design and implement programs that protect children before they become sick. The methods applied to the six study counties can be extended to the other 94 North Carolina counties and nationally. This modeling approach can be expanded to include exposure risk indices for a variety of children's environmental health issues, including asthma triggers, allergens, pesticides, and other chemicals. GIS technology holds tremendous potential for revolutionizing how environmental health organizations conceive their agendas as well as how they design and implement both conventional programs and those associated with emerging issues.

REFERENCES AND NOTES

- Lanphear BP, Weitzman M, Eberly S. Racial differences in environmental exposures to lead. *Am J Public Health* 86:1416–1421 (1996).
- Norman EH, Bordley WC, Hertz-Picciotto I, Newton DA. Rural-urban blood lead differences in North Carolina children. *Pediatrics* 94:59–64 (1994).
- Sargent J, Brown MJ, Freeman J, Bailey A. Childhood lead poisoning in Massachusetts communities: its associations with sociodemographic and housing characteristics. *Am J Public Health* 85:528–532 (1995).
- Bailey A, Sargent J, Blake M. A tale of two counties: childhood lead poisoning, industrialization, and abatement in New England. *Economic Geography AAG Special Issue* 96–111 (1998).
- Griffith DA, Doyle PG, Wheeler DC, Johnson DL. A tale of two swaths: urban childhood blood-lead levels across Syracuse, New York. *Ann Assoc Am Geogr* 88:640–645 (1998).
- Guthrie W, Tucker R, Murphy E, England R, Stevenson E, Luckhardt J. Reassessment of lead exposure in New Jersey using GIS technology. *Environ Res* 59:318–325 (1992).
- Lanphear B, Byrd RS, Auinger P, Schaffer SJ. Community characteristics associated with elevated blood lead levels in children. *Pediatrics* 101(2):264–271 (1998).
- Reissman D, Staley F, Curtis GB, Kaufmann R. Use of geographic information system technology to aid health department decision making about childhood lead poisoning prevention activities. *Environ Health Perspect* 109(1):89–94 (2001).
- Sargent JD, Bailey A, Simon P, Blak M, Dalton MA. Census tract analysis of lead exposure in Rhode Island children. *Environ Res* 74:159–168 (1997).
- Bellinger DC. Interpreting the literature on lead and child development: the neglected role of the "experimental system." *Neurotoxicol Teratol* 17(3):201–212 (1995).
- Needleman HL, Schell A, Bellinger D, Leviton A, Allred E. The long-term effects of exposure to low doses of lead in childhood: an 11-year follow-up report. *N Engl J Med* 322(2):83–88 (1990).
- Needleman HL, Gatonis CA. Low-level lead exposure and the IQ of children: a meta-analysis of modern studies. *JAMA* 263(5):673–678 (1990).
- Needleman HL, Riess J, Tobin M, Biesecker G, Greenhouse J. Bone lead levels and delinquent behavior. *JAMA* 275(5):363–369 (1996).
- Schwartz J. Low-level lead exposure and children's IQ: a meta-analysis and search for a threshold. *Environ Res* 65:42–55 (1994).
- Schwartz J, Otto D. Lead and minor hearing impairment. *Arch Environ Health* 46(5):300–305 (1991).
- Thacker SB, Hoffman DA, Smith J, Steinberg K, Zack M. Effect of low-level body burdens of lead on the mental development of children: limitations of meta-analysis in a review of longitudinal data. *Arch Environ Health* 47(5):336–346 (1992).
- Lanphear B, Dietrich K, Auinger P, Cox C. Cognitive deficits associated with blood lead concentrations < 10 microg/dL in US children and adolescents. *Public Health Rep* 115(6):521–529 (2000).
- Lanphear B, Canfield R, Henderson C, Cory-Slechta D, Cox C. Environmental exposure to lead and children's intelligence at blood lead concentrations below 10 micrograms per deciliter. Presented at the Pediatric Academic Societies' 2001 Annual Meeting, 28 April–1 May 2001, Baltimore, MD.
- Centers for Disease Control and Prevention. Update: blood lead levels—United States, 1991–1994. *Morb Mortal Wkly Rep* 46(7):141–146 (1997).
- CDC. Preventing Lead Poisoning in Young Children: A Statement by the Centers for Disease Control. Atlanta, GA:Centers for Disease Control and Prevention, 1991.
- Shannon M. Etiology of childhood lead poisoning. In: *Lead Poisoning in Childhood* (Pueschel S, Linakis J, Anderson A, eds). Baltimore, MD:Paul H. Brookes Publishing Company, 1996:37–58.
- U.S. GAO. Lead Poisoning: Federal Health Care Programs Are Not Effectively Reaching At-Risk Children. Washington, DC:United States General Accounting Office, 1999.
- North Carolina Childhood Lead Poisoning Prevention Program, North Carolina Department of Environment and Natural Resources databases. Unpublished data.
- CDC. Screening Young Children for Lead Poisoning: Guidance for State and Local Public Health Officials. Atlanta, GA:Centers for Disease Control and Prevention, 1997.
- Mahaffey KR, Gartside PS, Glueck CJ. Blood lead levels and dietary calcium intake in 1- to 11-year-old children: the Second National Health and Nutrition Examination Survey, 1976 to 1980. *Pediatrics* 78:257–262 (1986).
- Vine MF, Degan D, Hanchette C. Geographic information systems: their use in environmental epidemiologic research. *Environ Health Perspect* 105(6):598–605 (1997).
- Wartenberg D. Screening for lead exposure using a geographic information system. *Environ Res* 59:310–317 (1992).
- Wilson County Tax Assessor, Wilson County, NC. Unpublished data.
- Geography Network 2001. TIGER Data. Redlands, CA:Environmental Sciences Research Institute. Available: <http://www.geographynetwork.com> [cited 3 August 2001].